

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-181998

(43)Date of publication of application : 21.07.1995

(51)Int.Cl. G10L 9/14

(21)Application number : 05-327898

(71)Applicant : SANYO ELECTRIC CO LTD

(22)Date of filing : 24.12.1993

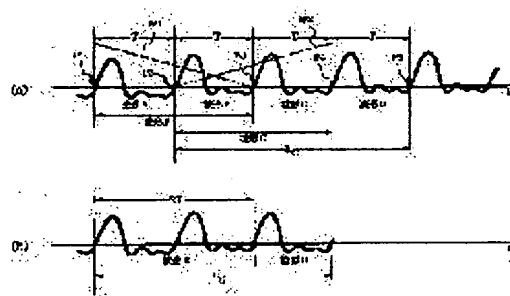
(72)Inventor : INOUE TATSUO  
SUGISHITA SHOZO

## (54) METHOD FOR COMPRESSING AND EXPANDING VOICE TIME-BASE

## (57)Abstract:

PURPOSE: To develop a voice compression and expansion system by which a voice waveform small in displacement and distortion can be obtained even when the compression/expansion ratio is close to 1.

CONSTITUTION: When a voice signal is compressed, voice signal waveforms F and G whose lengths are  $\frac{1}{2}$  times as large a pitch cycles and integral multiples are cut from both the position P1 of the current pointer of the inputted voice signal waveform and the position P2 which is one pitch cycle precedent to the pointer, and a couple of cut voice signal waveforms are weighted with proper W1 and W2 and added to generate one signal waveform E by compressing the two voice signal waveforms, and, the current pointer P1 is moved on the inputted voice signal waveform corresponding to compressibility to obtain a pointer P4, and similar operation is performed to obtain a compressed voice signal.



## LEGAL STATUS

[Date of request for examination] 22.04.1997

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 2951181

[Date of registration] 09.07.1999

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-181998

(43) 公開日 平成7年(1995)7月21日

(51) Int.Cl.<sup>6</sup>

G 1 0 L 9/14

識別記号

G

J

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数 2 O L (全 7 頁)

(21) 出願番号

特願平5-327898

(22) 出願日

平成5年(1993)12月24日

(71) 出願人 000001889

三洋電機株式会社

大阪府守口市京阪本通2丁目5番5号

(72) 発明者 井上 健生

大阪府守口市京阪本通2丁目18番地 三洋電機株式会社内

(72) 発明者 杉下 正蔵

大阪府守口市京阪本通2丁目18番地 三洋電機株式会社内

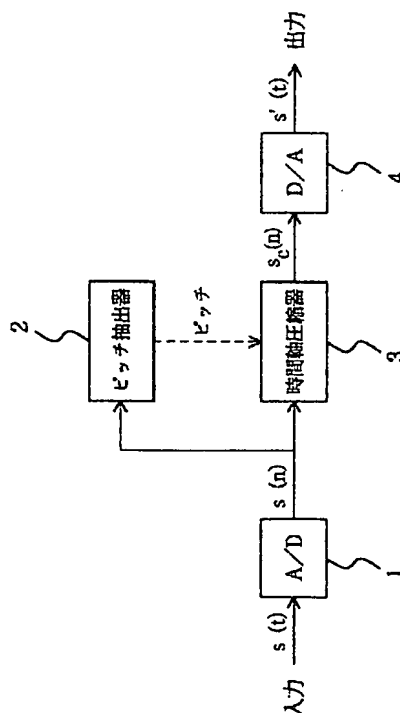
(74) 代理人 弁理士 岡田 敬

(54) 【発明の名称】 音声時間軸圧縮方法及び伸長方法

(57) 【要約】

【目的】 圧縮伸長率が1に近い時でもずれや歪みの小さい音声波形が得られる音声圧縮伸長方式を開発する。

【構成】 音声信号の圧縮時において、入力された音声信号波形の現在のポイントの位置P1と該ポイントに対して1ピッチ周期だけ進んだ位置P2との両方から夫々前記ピッチ周期の2倍以上且つ整数倍の長さの音声信号波形F、Gを切り出し、切り出された一対の音声信号波形を適当な重みW1、W2をつけて加算し、前記2つの音声信号波形を圧縮した1つの信号波形Eを生成し、次に前記現在ポイントP1を圧縮率に応じて前記入力された音声信号波形上で移動せしめてポイントP4として同様の操作を行い圧縮された音声信号を得る。



## 【特許請求の範囲】

【請求項 1】 音声信号の圧縮時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して 1 ピッチ周期だけ進んだ位置との両方から夫々前記ピッチ周期の 2 倍以上で且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記 2 つの音声信号波形を圧縮した 1 つの信号波形を生成し、次に前記現在ポイントを圧縮率に応じて前記入力された音声信号波形上で移動せしめて同様の操作を行い圧縮された音声信号を得ることを特徴とする音声時間軸圧縮方法。

【請求項 2】 音声信号の伸長時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して 1 ピッチ周期だけ遅れた位置との両方から夫々前記ピッチ周期の 2 倍以上で且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記 2 つの音声信号波形を合成した 1 つの信号波形を生成し、次に前記現在ポイントを伸長率に応じて前記入力された音声信号波形上で移動せしめて同様の操作を行い伸長された音声信号を得ることを特徴とする音声時間軸伸長方法。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】 本発明は、放送、映画、音楽において、音声処理する場合の発声速度を制御する音声の発声速度変換方法、及び音声 IC レコーダ等に用いられるデジタルの音声信号を主として符号化する音声の時間軸圧縮あるいは伸長方法に関する。

## 【0002】

【従来の技術】 従来デジタルの音声信号を時間軸圧縮あるいは伸長する方法として、IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-27, pp. 121-133, Apr. '79 "Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals" D. Malah 著に開示された TDHS (Time-Domain Harmonic Scaling) や「自己相関法による音声の時間軸での伸縮方式とその評価」電子通信学会春季講演予稿集 pp. 9-16, 1986. 5. 21 森田直孝、板倉文忠著に開示された P I C O L A (Pointer Interval Control Overlap and Add) 等が提唱されている。

【0003】 前記 TDHS では、音声波形に対して予めピッチ抽出器を用いて抽出されたピッチ周期を  $T$ 、圧縮率（あるいは伸長率）を  $r$  とするとき、図 7 (a) の波形図に示すように現在のポイント位置  $P_1$  と、これに対してピッチ周期  $T$  だけ進んだ時点  $P_2$  から夫々長さ  $N$

c、

## 【0004】

## 【数 1】

$$N_C = rT / (1 - r)$$

【0005】 の音声波形 A 及び B を切り出し、波形 A には 1 から 0 へ直線的に向かう重み（図中破線 F 1 で示す窓関数）を、一方の波形 B には逆に 0 から 1 へ直線的に向かう重み（図中破線 F 2 で示す窓関数）を付与して、両者を加え合わせることににより長さ  $N_C$  の波形 C を得る。次にポイントを前記入力波形（図 7 (a)）上で、

## 【0006】

## 【数 2】

$$N_C + T = T / (1 - r)$$

【0007】 だけ移動して、前記と同様にピッチ周期  $T$  の 2 つの波形を切り出し、これらに重みをつけて加え合わせることににより、次の長さ  $N_C$  の波形を得る。以後この操作を繰り返すことににより長さ  $N_C + T$  の入力波形から長さ  $N_C$  の出力波形が連続して作り出される。しかもこの長さ  $N_C$  の波形は長さ  $N_C + T$  の入力波形に対して圧縮率  $r$  を満足したものとなっている。

【0008】 一方、伸長の場合には図 8 (a) (b) の波形図に示すように現在のポイント位置  $P_1$  とこれに対してピッチ周期  $T$  だけ遅れた時点  $P_4$  から夫々長さ  $N$

s、

## 【0009】

## 【数 3】

$$N_S = rT / (r - 1)$$

【0010】 の音声波形 A 及び B を切り出し、波形 A には 0 から 1 へ直線的に向かう重み（図中破線 F 3 で示す窓関数）を、一方の波形 B には逆に 1 から 0 へ直線的に向かう重み（図中破線 F 4 で示す窓関数）を付与して、両者を加え合わせることににより長さ  $N_S$  の波形 C を得る。次にポイントを前記入力波形（図 8 (a)）上で、

## 【0011】

## 【数 4】

$$N_S - T = T / (r - 1)$$

【0012】 だけ移動して、前記と同様にピッチ周期  $T$  の 2 つの波形を切り出し、これらに重みを付けて加え合わせることににより、次の長さ  $N_S$  の波形を得る。以後この操作を繰り返すことににより長さ  $N_S - T$  の入力波形に対し伸長率  $r$  の出力波形が連続して得られる。しかもこの長さ  $N_S$  の波形は長さ  $N_S - T$  の入力波形に対して伸長率  $r$  を満足したものとなっている。

【0013】 ところでかかる TDHS を時間軸の圧縮あるいは伸長に適用すると、上述のように圧縮率あるいは伸長率  $r$  が 1 に近い場合、波形長  $N_C$  や  $N_S$  がピッチ周期  $T$  よりも大きくなり過ぎてしまい、加え合わせる音声波形内でピッチ周期の揺らぎによる波形のずれが生じ、歪みの原因になるという問題点があった。

## 【0014】

【発明が解決しようとする課題】 本発明は、上記のよう

にTDHSを時間軸の圧縮あるいは伸長に適用すると圧縮率（あるいは伸長率） $r$ が1に近い場合、 $N_c$ や $N_s$ が $T$ より大きく加え合わせる音声波形内でピッチ周期の揺らぎによる波形のずれが生じ、歪みの原因となるという従来の問題点に鑑み、圧縮率（あるいは伸長率）が1に近い時でもずれや歪みの小さい音声波形が得られる音声圧縮及び伸長方式を開発することを目的とする。

【0015】

【課題を解決するための手段】本発明は、音声信号の圧縮時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して1ピッチ周期だけ進んだ位置との両方から夫々前記ピッチ周期の2倍以上且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記2つの音声信号波形を圧縮した1つの信号波形を生成し、次に前記現在ポイントを圧縮率に応じて前記入力された音声信号波形上あるいは生成された音声信号上で移動せしめて同様の操作を行い圧縮された音声信号を得るものである。

【0016】また本発明は、音声信号の伸長時において、入力された音声信号波形の現在のポイントの位置と該ポイントに対して1ピッチ周期だけ遅れた位置との両方から夫々前記ピッチ周期の2倍以上且つ整数倍の長さの音声信号波形を切り出し、切り出された一対の音声信号波形を適当な重みをつけて加算し、前記2つの音声信号波形を合成した1つの信号波形を生成し、次に前記現在ポイントを伸長率に応じて前記入力された音声信号波形上あるいは生成された音声信号上で移動せしめて同様の操作を行い伸長された音声信号を得るものである。

【0017】

【作用】上記構成によると、加え合わせる音声波形の長さは、圧縮率あるいは伸長率に関わり無くピッチ周期の2倍以上且つ整数倍の長さであり、ポイントの移動量によって圧縮率あるいは伸長率の調整が行われるため、常に最も似通った波形同士が加え合わされ、圧縮率あるいは伸長率が1に近づくに従って得られた音声信号の歪みは小さくなる。

【0018】

【実施例】以下本発明の音声時間軸圧縮方法及び伸長方法をその一実施例について図面に基づいて詳細に説明する。

【圧縮方法】対応図は図1～図3

図1は当該圧縮方法を実現する回路ブロック図を示す。

【0019】図1においてアナログ音声信号 $s(t)$ はA/D変換器1によりデジタル音声データ $s(n)$ に変換される。次にピッチ抽出器2によりデジタル音声データ $s(n)$ のピッチ周期 $T$ が抽出され、抽出されたピッチ情報を用いて時間軸圧縮器3により音声データ $s(n)$ の時間軸圧縮が行われ、結果として圧縮されたデジタル音声信号 $s_c(n)$ を得る。

【0020】そして前記ピッチ周期情報 $T$ と圧縮された音声信号 $s_c(n)$ をD/A変換器4に送り、ここでアナログ音声信号 $s'(t)$ に変換して出力する。

【0021】以下に前記図1における時間軸圧縮器3の詳細な働きを図2(a)(b)及び図3(a)(b)を用いて説明する。

【0022】前記時間軸圧縮器3では図2(a)に示すように、夫々現在のポイント $P1$ と、これに対してピッチ周期 $T$ だけ進んだ時点 $P2$ から、夫々長さ $2T$ の音声波形 $F$ （音声波形 $A$ ＋音声波形 $B$ 、即ちポイント $P1 \sim P3$ の区間）及び音声波形 $G$ （音声波形 $B$ ＋音声波形 $C$ 、即ちポイント $P2 \sim P4$ の区間）を切り出し、波形 $F$ には1から0へ向かう重み（窓関数 $W1$ ）をつけ、波形 $G$ には逆に0から1へ向かう重み（窓関数 $W2$ ）をつけてこれら両波形を加え合わせ、長さ $2T$ の音声波形 $E$ を得る（図2(b)参照）。なお、ここでの重み $W1$ 、 $W2$ は音声波形 $E$ の前後の接続点での波形の連続性を保つために設けたものである。

【0023】次に図2の $L_c$ を、

【0024】

【数5】

$$L_c = rT / (1 - r)$$

【0025】とするとき、 $L_c \geq 2T$ のときには前記図2(a)の入力音声上で $T + L_c$ だけポイントを移動させてポイント $P5$ として上記と同様の操作を行う。この時の出力は前記波形 $E$ とこれに続く $P4$ から $P5$ までの長さ $L_c - 2T$ の入力波形をそのまま出力することになる。

【0026】一方 $L_c < 2T$ のときは図3(a)に示す入力音声上でポイント $P1$ から $T + L_c$ だけ右にポイントを移動させてポイント $P6$ として図2(a)と同様の操作を行う。この時の出力は波形 $E$ の前半の長さ $L_c$ の部分のみとなり、残りの $2T - L_c$ の部分は次の処理に再び用いられるため入力側へ戻される。

【0027】以上の操作によって長さ $T + L_c$ の入力波形から長さ $L_c$ の出力波形が作られ、これは音声信号の圧縮率 $r$ を満足するものとなっている。なお、図2

(a)(b)は圧縮率 $r = 3/4$ 、図3(a)(b)は圧縮率 $r = 3/5$ の場合を夫々示しており、双方とも圧縮波形を切り出す波形長はピッチ周期の2倍の場合を示している。

【0028】もちろん切り出す波形は2倍以上の整数倍であれば良くこの実施例に限られるものではないことは言うまでもない。

【伸長方法】対応図は図4～図6

図4は当該伸長方法を実現する回路ブロック図を示す。

【0029】図4においてアナログ音声信号 $s(t)$ はA/D変換器1によりデジタル音声データ $s(n)$ に変換される。次にピッチ抽出器2によりデジタル音声デー

タ  $s(n)$  のピッチ周期  $T$  が抽出され、抽出されたピッチ情報を用いて時間軸伸長器 5 により音声データ  $s(n)$  の時間軸伸長が行われ、結果として伸長されたデジタル音声信号  $s_e(n)$  を得る。

【0030】そして前記ピッチ周期情報  $T$  と伸長された音声信号  $s_e(n)$  を D/A 変換器 4 に送り、ここでアナログ音声信号  $s'(t)$  に変換して出力する。

【0031】以下に前記図 4 における時間軸伸長器 5 の詳細な働きを図 5 (a) (b) 及び図 6 (a) (b) を用いて説明する。

【0032】前記時間軸伸長器 5 では図 5 (a) に示すように、現在のポイント  $P1$  と、このポイント  $P1$  に対してピッチ周期  $T$  だけ遅れた時点  $P4$  から夫々長さ  $2T$  の音声波形  $F$  (波形  $A$  + 波形  $B$ 、即ちポイント  $P4 \sim P2$  の区間) 及び音声波形  $G$  (波形  $B$  + 波形  $C$ 、即ちポイント  $P1 \sim P3$  の区間) を切り出し、波形  $F$  には 0 から 1 へ向かう重み  $W3$  を、波形  $G$  には 1 から 0 へ向かう重み  $W4$  をつけて加え合わせ、図 5 (b) に示す長さ  $2T$  の音声波形  $H$  を作成する。この時の重み  $W3$ 、 $W4$  は前記時間軸圧縮器 3 の場合と同様に音声波形  $H$  の前後の接続点での連続性を保つために設けるものである。

【0033】そして  $L_s$  を

【0034】

【数 6】

$$L_s = T / (r - 1)$$

【0035】とするとき、 $L_s \geq T$  の時は前記図 5 (a) の入力音声上で  $L_s$  だけ移動させてポイント  $P3$  として同様の操作を行う。この時の出力は波形  $H$  とこれに続くポイント  $P2 \sim P3$  までの長さ  $L_s - T$  の入力波形 (波形  $C$ ) をそのまま出力する。一方、 $L_s < T$  のときは図 6 (a) の入力音声上でポイント  $P1$  から  $L_s$  だけポイントを移動させて、ポイント  $P5$  として上記と同様の操作を行う。この時の出力は、波形  $H$  の前半の長さ  $T + L_s$  の部分のみとなり、残りの  $T - L_s$  の部分は次の処理に再び用いられるため、入力側に戻される。

【0036】以上の操作によって長さ  $L_s$  の入力波形から長さ  $T + L_s$  の出力波形が作られており、伸長率  $r$  が満足されていることが解る。なお、図 5 (a) (b) は

伸長率  $r = 3/2$  の場合を示しており、図 6 (a)

(b) は伸長率  $r = 3$  の場合で共に伸長波形を作る時に切り出す波形長はピッチ周期の 2 倍の場合を示している。もちろん切り出す波形は 2 倍以上の整数倍であれば良くこの実施例に限られるものではないことは言うまでもない。

【0037】

【発明の効果】本発明の時間軸圧縮方法及び伸長方法によると、圧縮時または伸長時に加え合わせる音声波形の長さは圧縮率あるいは伸長率に関わり無くピッチ周期の 2 倍以上の整数倍の長さであり、ポイントの移動量によって圧縮あるいは伸長の調整が行われるため、ピッチ周期の揺らぎによる波形のずれが抑えられ、結果的に圧縮あるいは伸長時の音声波形の歪みが小さくなる効果が期待できる。

【0038】また従来の方と比較して圧縮率あるいは伸長率が 1 に近づくに従って歪みが小さくなる効果も期待できる。

【図面の簡単な説明】

【図 1】本発明の時間軸圧縮方法を実現するブロック回路図である。

【図 2】本発明方法による圧縮時 ( $L_c \geq 2T$ ) の入力波形 (a) 及び出力波形 (b) を示す図である。

【図 3】本発明方法による圧縮時 ( $L_c < 2T$ ) の入力波形 (a) 及び出力波形 (b) を示す図である。

【図 4】本発明の時間軸伸長方法を実現するブロック回路図である。

【図 5】本発明方法による伸長時 ( $L_s \geq T$ ) の入力波形 (a) 及び出力波形 (b) を示す図である。

【図 6】本発明方法による伸長時 ( $L_s < T$ ) の入力波形 (a) 及び出力波形 (b) を示す図である。

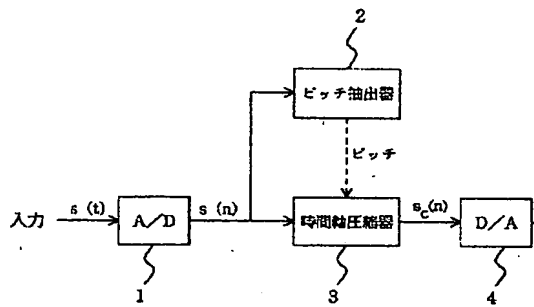
【図 7】図 2、図 3 に対応する従来方法の図である。

【図 8】図 5、図 6 に対応する従来方法の図である。

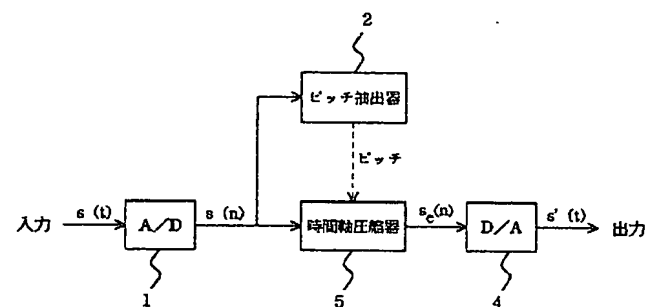
【符号の説明】

1	A/D 変換器
2	ピッチ抽出器
3	時間軸圧縮器
4	D/A 変換器
5	時間軸伸長器

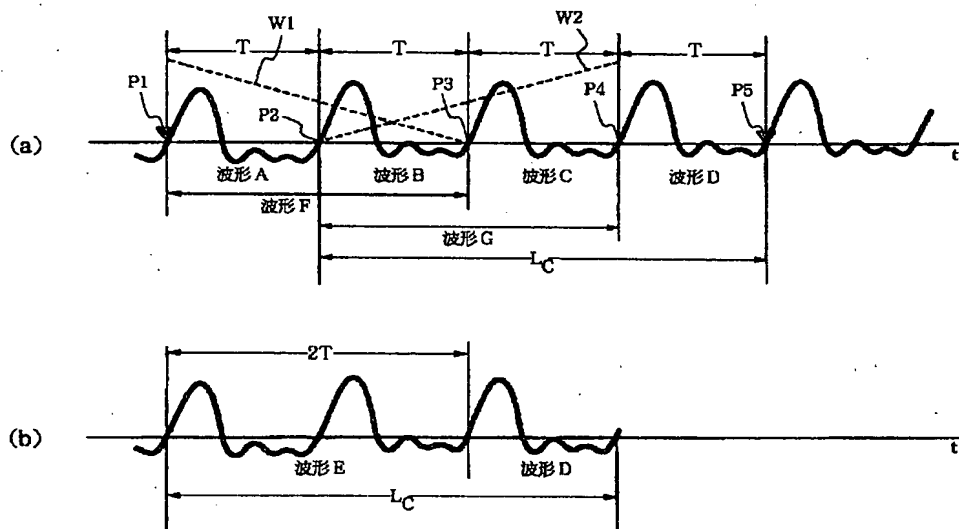
【図 1】



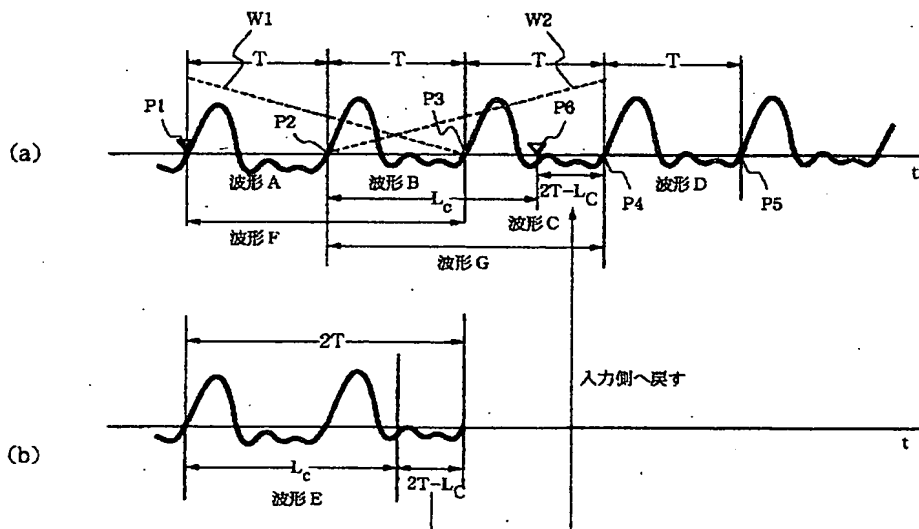
【図 4】



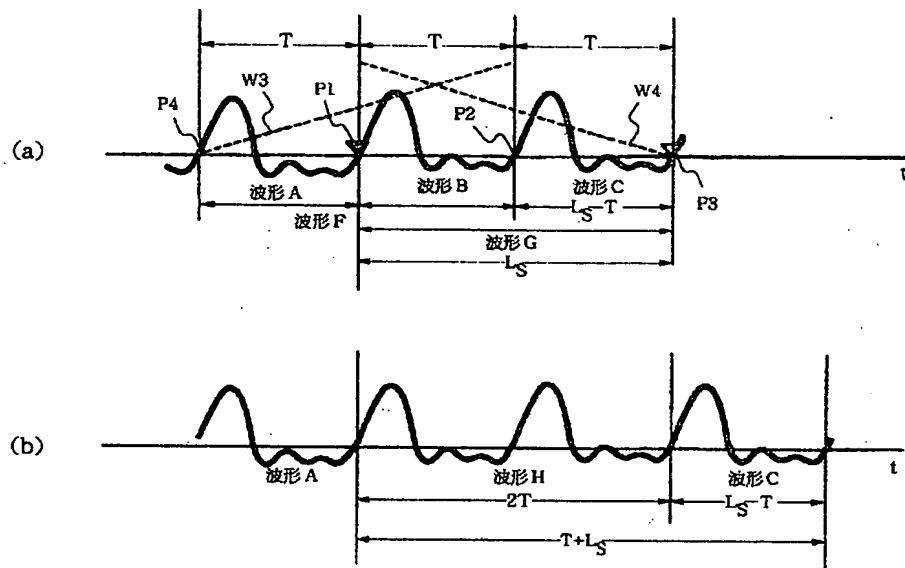
【図 2】



【図 3】



【図 5】



【図 6】

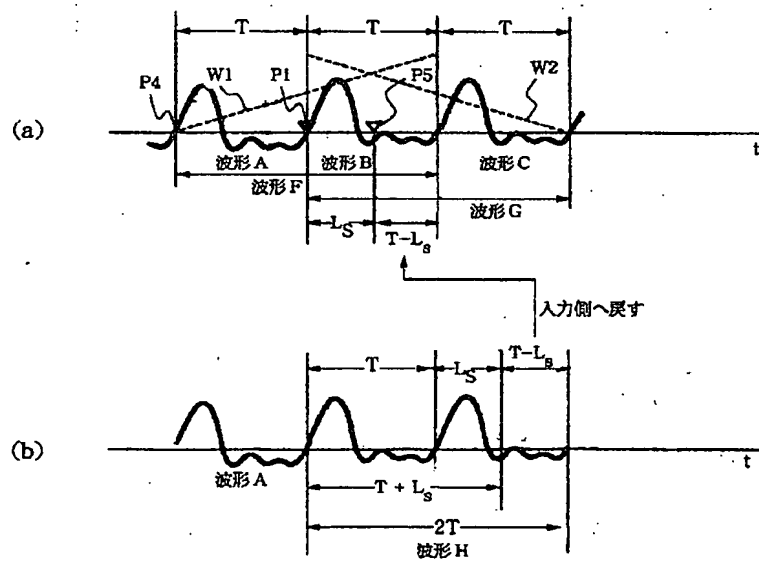


Figure 1 consists of two parts, (a) and (b), illustrating a signal processing method. Part (a) shows a continuous signal waveform over time  $t$ . The waveform is divided into two segments, '波形 A' and '波形 B', by a dashed vertical line. The total duration of the segments is  $T$ . The duration of each segment is  $N_c$ . Points  $P1$ ,  $F1$ ,  $P2$ , and  $F2$  are marked on the waveform. Part (b) shows a single segment '波形 C' with duration  $N_c$ .